

ARMY RESEARCH LABORATORY



Acoustic Feature Extraction for a Neural Network Classifier

Mark C. Wellman, Nassy Srour, and David B. Hillis

ARL-TR-1166

January 1997

19970131 036

DTIC QUALITY INSPECTED 2

Approved for public release; distribution unlimited.

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory

Adelphi, MD 20783-1197

ARL-TR-1166

January 1997

Acoustic Feature Extraction for a Neural Network Classifier

Mark C. Wellman, Nassy Srour, and David B. Hillis

Sensors and Electron Devices Directorate

Abstract

Artificial neural networks can perform reliable classification of ground vehicles based solely on their acoustic signatures, if robust features can be identified. We present feature extraction and classification results using simple power spectrum estimates, harmonic line association, and principal component analysis. Algorithm implementation and performance analysis of each feature extraction method are discussed. Also given are preliminary evaluation results of a VLSI (very-large-scale integration) device dedicated to neural network implementation.

Contents

| | |
|---|-----------|
| 1. Introduction | 1 |
| 1.1 <i>Feature Extraction</i> | 1 |
| 1.2 <i>Feature Extraction Techniques</i> | 1 |
| 1.3 <i>Artificial Neural Networks</i> | 2 |
| 2. Procedure | 4 |
| 2.1 <i>Data Collection</i> | 4 |
| 2.2 <i>Feature Extraction Methods</i> | 5 |
| 2.3 <i>Backpropagation Neural Network</i> | 6 |
| 3. Results | 8 |
| 3.1 <i>Confusion Matrices</i> | 8 |
| 3.2 <i>Results using CNAPS</i> | 9 |
| 4. Conclusions | 10 |
| 5. Future Considerations | 11 |
| References | 13 |
| Distribution | 15 |
| Report Documentation Page | 20 |

Figures

| | |
|---|---|
| 1. General neuron structure | 3 |
| 2. Neural network architecture | 3 |
| 3. Field sensor and processing architecture | 5 |

Tables

| | |
|---|---|
| 1. Classes of vehicles | 5 |
| 2. Testing results for trained BPNN | 8 |

1. Introduction

It should be possible to use artificial neural networks (ANNs) to classify tracked and wheeled vehicles solely based on their acoustic signatures. The main problem faced in classification is the selection of proper feature vectors that will be stable and class specific. Acoustic signatures are typically nonstationary [1,2] and are often corrupted by propagation effects, noise, and interference from the environment [1,3,4]. A robust feature extraction technique must be, to some degree, tolerant of these issues in order to be reliable. We have investigated three feature extraction techniques: simple power spectrum estimates (PSEs), harmonic line association (HLA) techniques [4], and principal component analysis (PCA) [5–7]. Algorithm implementation and performance analysis for these techniques are discussed and compared. Also given are preliminary evaluation results of a VLSI (very-large-scale integration) chip dedicated to neural networks.

1.1 Feature Extraction

Fundamentally, feature extraction and selection involve choosing those features of a class of patterns (whether waveforms, images, or geometric shapes) that will maintain class separability under the constraint of some criterion function. Feature extraction is a mapping of the original n -dimensional measurements into an m -dimensional feature space ($n > m$). In theory, the Bayes error [5,7,8] is the optimum measure of a feature's effectiveness, but it is difficult to obtain. One would need to perform nonparametric density estimation [5,7,8], a very time-consuming task, to obtain the posterior probabilities and in turn the Bayes error. Often in practice, feature extraction for representation is different from that for classification: features used for representation can be suboptimal for classification since they are not based on class separability [5]. The criterion used frequently for systematic feature extraction (Fukunaga's separability criterion) is based on a family of scatter matrices that can measure the class separability and generate optimal transformation matrices. This criterion can be applied by ANNs with the correct training algorithm [6,9–12]. The main problem is in its implementation under current system constraints; therefore, we use simpler analytical tools to evaluate the feature space.

1.2 Feature Extraction Techniques

The spectral characteristics of vehicle noise are distinctive: their acoustic signatures are dominated by narrow-band spectral peaks, since the physical process producing these sounds (engine firing rate and track slap) is periodic [3]. Spectral methods are amenable to calculation because of their simplicity and the existence of fast algorithms. These spectral lines, the first feature space considered, should present a good feature vector for classification and have in fact been used in the past for classification based on simple clustering techniques [4], for hierarchical clustering, and as inputs to an ANN [13]. The spectral peaks are typically bandlimited between 0 and 400 Hz, but peak components occur between 10 and 120 Hz.

A second feature space based on HLA [4] allows one to reduce the feature space considerably and should not appreciably reduce the separability of the various classes of vehicles [5].

A third feature space that holds promise is the principal components. PCA is a well-established technique in feature selection for both representation and classification [5,6,9]. PCA has a high degree of energy compaction: it basically transforms the original space into an uncorrelated space, thus reducing the dimension of the feature space. PCA is the brother of the Karhunen-Loeve transform, which is known to be the optimal transform method for signal representation [14,15]. Principal components are derived by the following set of relationships. Let

$$\mathbf{u}(n) = [u(t), u(t+1), \dots, u(t+N-1)]^T \quad (1)$$

be an $N \times 1$ random input vector and assume zero mean without loss of generality. Let \mathbf{R} be the $N \times N$ correlation matrix of the data with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_N$. The k principal components are defined by the linear transformation

$$\mathbf{C}(n) = \Phi^T \mathbf{u}(n) , \quad (2)$$

where $\mathbf{C}(n) = [c_1, c_2, \dots, c_k]^T$ is a $k \times 1$ principal component matrix, and Φ is an $N \times k$ matrix, with columns corresponding to k eigenvectors for the k largest eigenvalues of \mathbf{R} . We choose the value k arbitrarily, where k is significantly less than N , thus reducing the dimensionality of the original input space. The correlation matrix of the newly formed principal components is a diagonal matrix; thus the principal components are uncorrelated [6,19]. The principal components are optimal in a mean square sense and have removed redundancy associated with the original measurement. A motivation for using principal components is that data that exhibit a high degree of correlation from sample to sample may allow fast algorithms to implement PCA. Also, several researchers in neural networks [6,9–11,15,16] have derived learning algorithms to implement PCA. Finally, work in perfect reconstruction filter banks [17] leads one to believe that it may be possible to employ PCA in “real time.”

1.3 Artificial Neural Networks

ANNs are currently in use by ARL for classification of ground and air targets. In target classification, the ANN can not only aid in providing information about the target class, but also give a measure of one’s confidence in the decision. ANNs derive their computational power from their parallel distributed structure and ability to learn. Because neurons are basically nonlinear devices, the ANN will be nonlinear; nonlinearity is a very important property in light of the input signal structure, which is nonstationary and perhaps nonlinear.

The backpropagation ANN derives its name from the error-correction rule used in its training. Basically, the error backpropagation consists of two passes through the network. The forward pass takes the input vector and computes an activity pattern that propagates through the network, layer

by layer. During the forward pass, all the synaptic weights remain fixed. In the backward pass, the synaptic weights are adjusted by an error correction rule that is fundamentally the same for both hidden neurons and output neurons and is based on stochastic gradient descent [9,19]. Each neuron has the general structure shown in figure 1. The overall architecture for the network is shown in figure 2.

From figure 1, the governing expression for the output of a single neuron is the summation of weighted inputs:

$$v_j(n) = \sum_{i=0}^p w_{ji}(n) x_i(n) \quad (3)$$

where

$$y_j(n) = \varphi_j(v_j(n)) \quad (4)$$

is the output of the j th neuron.

Figure 1. General neuron structure.

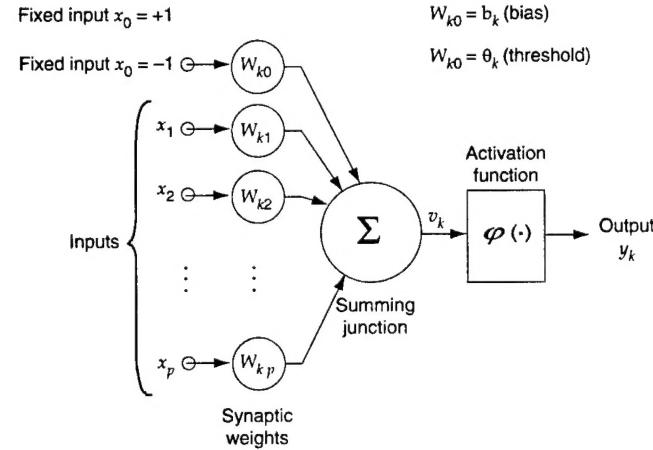
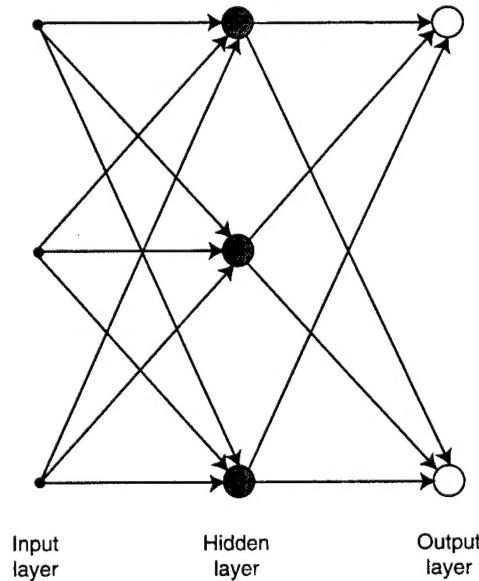


Figure 2. Neural network architecture.



Here $\phi_j(n)$ is the sigmoidal activation function given by

$$y_j = \frac{1}{1 + e^{-v_j}} , \quad (5)$$

an important approximation to hardlimiting. The sigmoidal activation function is differentiable, which facilitates the weight update given by the delta rule general expression

$$\Delta w_{ji}(n) = \eta \delta_j(n) y_i(n) ; \quad (6)$$

here η is the learning rate parameter, $\delta_j(n)$ is the local gradient (which depends on whether neuron j is in the output layer or the hidden layer), and $y_i(n)$ is the input to the j th neuron. The local gradient points to the required changes in the synaptic weights; in the output layer, the local gradient has the form

$$\delta_j(n) = e_j(n) \phi'_j(v_j(n)) , \quad (7)$$

with the error signal $e_j(n)$ given by

$$e_j(n) = d_j(n) - y_j(n) \quad (8)$$

and $d_j(n)$ is the desired signal.

The picture is more complex for the weight updates in the hidden layers; here the local gradient is dependent on all the errors associated with the neurons in the output layer when only one layer is hidden.

The local gradient for a hidden layer is given by

$$\delta_j(n) = \phi'_j(v_j(n)) \sum_k \delta_k(n) W_{kj}(n) \quad (9)$$

with $\delta_k(n)$ derived from the error signals associated with the k output neurons connected to the j th hidden neuron. These equations represent the general backpropagation algorithm and do not include the refinements available for a more robust network.

2. Procedure

2.1 Data Collection

RNADS (Remote Netted Acoustic Detection System) [1], a remote sensor architecture, was used to gather acoustic data from ground vehicles at Grayling, Michigan, and Aberdeen Proving Ground, Maryland. The vehicles included three tracked (class 0, 1, and 3) and one wheeled vehicle (class 2), all powered by 12-cylinder diesel engines (see table 1). The remote sensor consists of an 8-ft-diameter circular array of Knowles BL1994 ceramic microphones, with six microphones placed along the perimeter and a seventh microphone at the center of the array. This array baseline provides good directivity at low frequencies. Figure 3 shows the RNADS sensor and processing architecture.

Table 1. Classes of vehicles.

| Class | Vehicle |
|-------|--|
| 0 | 12-cylinder diesel, tracked vehicle |
| 1 | 12-cylinder diesel, tracked vehicle |
| 2 | 12-cylinder diesel engine, heavy wheeled vehicle |
| 3 | 12-cylinder diesel, tracked vehicle |

The acoustic signals were preamplified with a selectable gain of 40 and 60 dB and passed to a ruggedized personal computer (PC) and a digital audio tape (DAT) recorder. The DAT recorder sampled the acoustic signatures at a 2-kHz rate, well above the Nyquist rate. Within the PC, acoustic signals are anti-aliased with a lowpass filter, fed to 16-bit analog-to-digital converters, and further processed with a pair of commercially available digital signal processing boards for real-time applications.

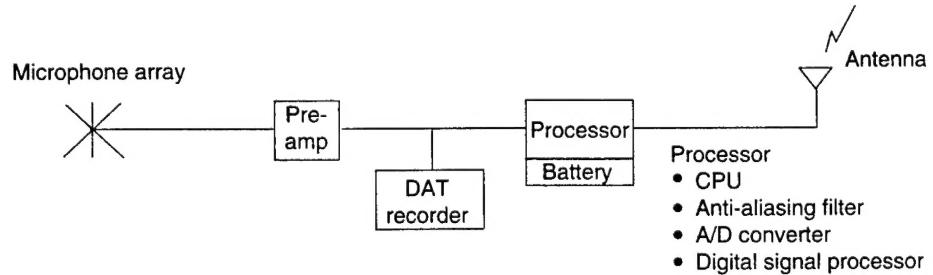
2.2 Feature Extraction Methods

We generated PSEs for each 1-s interval of data using Hanning windowed short-time Fourier transforms according to the Welch method [18]. We used the first 200 frequency bins derived from the power spectrum in the 1 to 200 Hz range for classification in the ANN.

A second technique used was selecting only those peaks that were "harmonically related." An HLA was developed by Robertson and Weber [4] to create harmonic line sets for each second of data samples. This algorithm takes the strongest peak P in the frequency peak set subject to the constraint $f_{fund} \in [8,20]$ Hz, assumes that this peak is some k th harmonic line of the fundamental frequency, and then calculates the total signal strength in that HLA set. The integer value k that gives the maximum signal strength is assumed to be the correct harmonic line number, and a total of 11 harmonic lines are retained as the feature vector. This technique has the advantage of normalizing the feature vector, since the feature is based on harmonic line number and not a function of frequency.

To calculate the principal components, we downsampled the data to 512 Hz and then divided them into $512/N$ subblocks of N samples ($N = 64$ or 128) for each data snapshot. The data were then used to generate a set of instantaneous autocorrelation matrix estimates (see eq (10) to (12)) [19]. Although subblock sizes of 64 and 128 samples were used in generating the correlation matrix estimates, we report only the results with 64-sample

Figure 3. Field sensor and processing architecture.



subblock sizes. The 128-sample subblocks gave similar classification results in preliminary training and testing of the neural network. The correlation matrix estimates were used for eigenanalysis; 11 eigenvectors associated with the largest 11 eigenvalues were used to transform the original data vector (64 samples) from each subblock to produce the principal components for that subblock. The principal components generated for each subblock were then averaged to produce an averaged principal component feature vector over the 512-sample block (see eq (13)). The entire procedure was repeated, shifting 256 samples (50-percent overlap) and forming a new 512-sample block for processing. We wanted to compare the performance of the PCA and HLA feature space in the classification scheme, so only 11 PCA features were retained; it was also necessary to generate one PCA feature vector per second of data sampled. The estimation process for a correlation matrix is based on the following data matrix formulation:

$$\mathbf{A}^H = [\mathbf{u}(M), \mathbf{u}(M+1), \dots, \mathbf{u}(N)] , \quad (10)$$

with the matrix $\mathbf{u}(i)$ given by

$$\mathbf{u}(i) = [u(i), u(i-1), \dots, u(i-M+1)]^T , \quad (11)$$

with the indices i falling in the range $[M, N]$.

Therefore, the data matrix \mathbf{A}^H is an M by $N - M + 1$ rectangular Toeplitz matrix.

Then the estimation of matrix \mathbf{R} is performed by

$$\mathbf{R} = \frac{1}{2(N-M)} \mathbf{A}^H \mathbf{A} , \quad (12)$$

and here the estimate \mathbf{R} will be an M by M matrix. The values used for actual processing of the PCA feature vectors were $M = 32, 64$ and $N = 64, 128$. The estimate of the principal components for the 512-sample block were generated from

$$\tilde{\mathbf{C}}_{block} = \frac{1}{l} \sum_{i=1}^l \Phi_i \mathbf{u}_i , \quad (13)$$

with $l = 512/N$ and the terms within the summation derived from equation (2).

2.3 Backpropagation Neural Network

The backpropagation neural network (BPNN) was trained by repeated presentation of examples of a particular input/output class with a subsequent adjustment to the synaptic weights based on the difference between the desired and the actual output. This process is repeated until the user set exit criterion is met for termination of the training procedure. Three different statistics can be used as exit criteria to terminate the training of the BPNN. The first is the number of epochs, which is a constant number of iterations assigned for training before training begins. The second is based on the mean square error (MSE), which is a general measure of the performance of a given neural network model for a given data set. The third exit

criterion in the BPNN is based on the *R*-squared statistic, which is the proportion variability in the target data set based on the input variables.

We used the epoch training as our exit criterion with a value of 1000 iterations. The learning rate parameter was set to 0.0005 and was automatically adjusted downward by an annealing divisor of 1.1. This adaptation of the learning rate allows fine-grain adjustments during the training. The maximum initial weights of the network were set to 0.01, and a random number generator was used to initialize these weights so that the network will avoid starting near a local minimum or an undesirable initial weight position. Further refinements to the learning rate were accomplished through an interlayer multiplier, which only affected the learning rate of the hidden neuron. The interlayer multiplier will cause the hidden nodes to be more sensitive to learning and thus improve the speed of learning.

Finally, smoothing was incorporated in the rate of learning for the BPNN. Smoothing can be highly beneficial to the learning behavior of the neural network [9]; it allows control of the weight adjustment based on the past values of gradient descent and can prevent the training process from terminating in a shallow local minimum. The greater the smoothing factor, the greater the influence of past adjustments and the smoother the migration of weights. A smoothing constant value of 0.9 was used in our neural network, which means that 90 percent of the weight adjustment is governed by the average of the past directions of gradient descent, and 10-percent by the current direction of gradient descent. This is the default smoothing constant in the Database Mining Workstation [13], a commercially available software package for unearthing and evaluating data characteristics using BPNNs. The data sets were divided into training and testing blocks for this purpose. Training sample sets were composed of 75, 67, and 50 percent of the data set.

The BPNN classifier was used to calculate the percentage of correct identification of ground vehicles, and in some cases the confidence levels were also generated. A confusion matrix was calculated that provides the percentage of correct identification (C_{ID}) for each class of ground vehicles based on

$$C_{ID} = \frac{N_{P_{CID}}}{K} . \quad (14)$$

Here $N_{P_{CID}}$ is the total number of correct predicted values and is the total number of observations. Confidence levels for the classification of each target were calculated by

$$C_1 = \frac{1}{K} \sum_{i=1}^K \left(\sum_{j=1}^L \frac{P_{CID_i} - P_{FID_j}}{2(L-1)} \right) , \quad (15)$$

where L is the total number of output classes, P_{CID_i} is the predicted value of correct identification for class i , and P_{FID_j} is the predicted value of false identification for class j with respect to class i .

Thus P_{CID_i} is the output for the neural network output node dedicated to a particular class, and P_{FID_j} is the output for the other output nodes.

3. Results

3.1 Confusion Matrices

Table 2 shows the confusion matrices for testing the PSE, HLA, and PCA features on the trained BPNN for each feature space. The numbers represent percentage of correct identification. The BPNN used had one input layer of 11 or 200 input nodes, one hidden layer of 15 nodes, and one output layer of 4 nodes. Table 1 (sect. 2) shows the general class characteristics of classes 0, 1, 2, and 3.

The scores in table 2 are representative of several trials for each BPNN and type of feature vector investigated; the values are rounded to the nearest integer. The rows do not sum to 100 percent because of roundoff error.

Table 2. Testing results for trained BPNN.

| Feature space | | Confusion matrices according to percentage trained/tested | | | | | | | | | | | | | |
|---------------|------------------------|---|----|----|------------------------|--------|------------|----|------------------------|-------|--------|------------|----|----|----|
| | | 75/25 | | | | 67/33 | | | | 50/50 | | | | | |
| PSE | Actual | Net output | | | | Actual | Net output | | | | Actual | Net output | | | |
| | | 0 | 1 | 2 | 3 | | 0 | 1 | 2 | 3 | | 0 | 1 | 2 | 3 |
| PSE | 0 | 93 | 0 | 3 | 3 | 0 | 90 | 0 | 4 | 4 | 0 | 93 | 0 | 0 | 4 |
| | 1 | 0 | 96 | 2 | 1 | 1 | 0 | 98 | 0 | 1 | 1 | 0 | 93 | 2 | 2 |
| | 2 | 3 | 0 | 96 | 0 | 2 | 4 | 0 | 95 | 0 | 2 | 8 | 0 | 89 | 2 |
| | 3 | 14 | 2 | 0 | 82 | 3 | 15 | 3 | 2 | 79 | 3 | 13 | 4 | 0 | 82 |
| | Confidence level = 81% | | | | Confidence level = 80% | | | | Confidence level = 78% | | | | | | |
| HLA | Actual | Net output | | | | Actual | Net output | | | | Actual | Net output | | | |
| | 0 | 0 | 1 | 2 | 3 | 0 | 0 | 1 | 2 | 3 | 0 | 0 | 1 | 2 | 3 |
| | 0 | 88 | 1 | 0 | 9 | 0 | 89 | 1 | 0 | 8 | 0 | 92 | 1 | 1 | 4 |
| | 1 | 0 | 93 | 6 | 0 | 1 | 0 | 94 | 6 | 0 | 1 | 0 | 89 | 10 | 0 |
| | 2 | 0 | 2 | 97 | 0 | 2 | 0 | 2 | 97 | 0 | 2 | 0 | 3 | 96 | 0 |
| | 3 | 20 | 0 | 12 | 67 | 3 | 20 | 1 | 11 | 66 | 3 | 27 | 3 | 10 | 58 |
| | Confidence level = 75% | | | | Confidence level = 74% | | | | Confidence level = 71% | | | | | | |
| PCA | Actual | Net output | | | | Actual | Net output | | | | Actual | Net output | | | |
| | 0 | 0 | 1 | 2 | 3 | 0 | 0 | 1 | 2 | 3 | 0 | 0 | 1 | 2 | 3 |
| | 0 | 98 | 0 | 0 | 0 | 0 | 95 | 0 | 2 | 2 | 0 | 92 | 1 | 3 | 3 |
| | 1 | 0 | 99 | 0 | 0 | 1 | 0 | 99 | 1 | 0 | 1 | 0 | 99 | 0 | 0 |
| | 2 | 6 | 0 | 90 | 2 | 2 | 6 | 1 | 84 | 7 | 2 | 7 | 0 | 86 | 6 |
| | 3 | 4 | 0 | 0 | 94 | 3 | 5 | 0 | 0 | 94 | 3 | 5 | 0 | 1 | 93 |
| | Confidence level = 88% | | | | Confidence level = 83% | | | | Confidence level = 81% | | | | | | |

3.2 Results using CNAPS

Further testing was performed with a commercial ANN software package known as "BrainMaker," developed by California Scientific Software [20], which was run on a general-purpose digital machine called CNAPS (Connected Network of Adaptive Processors) [21–23]. CNAPS, manufactured by Adaptive Solutions, is based on VLSI technology and is capable of high neural network performance. CNAPS is an SIMD (single instruction stream, multiple data stream) machine consisting of an array of 128 digital signal processors operating in parallel, significantly accelerating both the training and testing of ANNs. An 8-chip CNAPS system running at a mere 25 MHz, for example, can perform 12.8 billion multiply accumulates per second [9]. The best efficiencies are obtained with very large nets where up to 128 nodes in the same layer may be processed simultaneously, but smaller nets can gain some benefit as well.

The BrainMaker package running on the CNAPS hardware can train and run ANNs with exactly one input layer, one hidden layer, and one output layer using the standard backpropagation training algorithm. The user can set a training tolerance so that only those training examples with a root mean square (RMS) error above the tolerance will cause the weights to be updated. During testing, a user-supplied tolerance is used to determine the correctness of the net's answers. Test examples are scored as correct if the RMS error is below the tolerance, and they are scored as incorrect if the RMS error exceeds the tolerance. This is a somewhat conservative criterion, in that an example might have the highest activation in the correct output node, but still count as a misclassification if the RMS error were high.

For testing, a trial and error approach was used to find a good net configuration for the PCA data and for the HLA data. Training times were fixed at 2000 epochs for these tests. We determined the performance for each data set by averaging the results of 25 tests.

We used a random process to select 10 percent of the data set for testing and 90 percent for training. Five different training/testing set divisions were made, and five different training/testing cycles were performed for each division, for a total of 25 different tests. The percentage of correct classifications on the test set for each test was averaged to provide a single score. The average correct classification on the test sets was 90.8 percent for the HLA features and 96.8 percent for the PCA features. Training times were considerably less for the CNAPS card over software implementation: for example, for a neural network of 11 inputs, 11 hidden layers, and 4 output nodes, using epoch training of 4000 iterations and 516 feature vectors, the software trained in 13 minutes, whereas the CNAPS would perform the same training in 90 s.

4. Conclusions

The PCA features show a marked improvement over the HLA feature set for correct classification and confidence levels for all classes and a slight improvement over the PSE feature space. The most notable improvement was in the identification of class 3, where PCA feature extraction gave a far more robust and stable feature vector for this target class. Both of the other feature sets have difficulty discriminating this target class. The largest degree of misclassification for all feature vectors occurs between class 0 and 3, where as much as 27-percent misclassification occurs with the HLA features (see table 3). This result was also observed when hierarchical k -means clustering analysis was used to derive data clusters for the four classes (unpublished findings); again a great deal of crossover occurs for these two classes, with a lesser degree between class 2 and class 3. It is also interesting that even though PCA performs so well on the average, its performance in the classification of class 2 (a 12-cylinder diesel engine truck) is unexpectedly poor: this vehicle is very loud, with a characteristic signature, and has been classified to 100 percent using maximum likelihood methods in the past [3]. We would expect to see similar results with the PCA features, but this is not the case; perhaps the distinction is degraded by the fact that classes 0, 2, and 3 are closely related, and in this instance, the data were collected from the same environment. Further testing of trained neural networks with appropriate class data collected from other test sites should allow us to resolve this issue. We should be careful in considering the results for class 1 since it had a small representation in the training and testing: only one data file (albeit quite large) was used for this vehicle class. Class 1 data were also the only representative data collected at Aberdeen Proving Ground.

It is not surprising that the PSE feature space produced such a high degree of correct classification. The PSE results indicate that the narrowband features for each class are indeed highly class specific; a feature method that maintains some of the "brute force" frequency and amplitude resolution characteristic of PSE with lower dimensionality may be ideal. Despite the PSE's simplicity and performance, we expect that the classification results will drop for targets evaluated under different background environments, since the algorithm inherently has a high degree of sensitivity to environmental variables. Apparently, HLA features are lacking in some necessary narrowband components for a higher degree of correct classification.

The choice of using a simple backpropagation neural network classifier is supported not only by the results but also by theory: the backpropagation algorithm is not only simple in implementation but will closely approximate the Bayes error with increased training [9]. Preliminary results with the CNAPS card also support the notion that it will provide advantages to a fielded neural network classifier. When retraining is necessary, the CNAPS implementation will significantly enhance overall system robustness by its processing speed alone.

5. Future Considerations

For future work, the three feature spaces have to be evaluated for their complexity and real-time implementation. The downside to the PCA is that it is difficult to implement without resorting to the incorporation of a preprocessing feedforward neural network [6,9–12,16] or to periodic eigenanalysis of the acoustic data. The preprocessing neural network could be initialized in the field with scenario-based “eigenclusters” determined by the use of clustering techniques and derived for a set of classes that one would expect to encounter. These eigenclusters, which would be class specific (i.e., tracked versus wheeled), would be a one-dimensional application of Sirovich and Kirby’s “eigenpicture” method for classification [24]. Any subsequent retraining would be performed only when misclassification grows beyond some threshold. The instantaneous estimate of the correlation matrices by the simple matrix technique in this report is simply too time consuming when large data blocks are concerned, and thus direct procedures to calculate the principal components for each time block probably cannot be employed. Alternatively, we could average the autocorrelation matrix estimates over the entire 512-sample block, calculate the eigenvectors, and then transform each subblock to generate principal component estimates. Although this approach gives a considerable savings computationally and is less sensitive to signal to noise issues, it requires that we assume stationary signals over the sample block of interest. For nonstationary signals, the errors associated with this procedure may make the formation of principal components irrelevant. This technique will be investigated further because it may prove promising in “PCA-like” feature extraction.

Although the simple PSE feature space is readily derived and can be used rapidly for identification, preliminary results have shown that it is sensitive to the environment, and misclassification can grow substantially. Also, it is very time-consuming in the training stage; without the implementation of VLSI, it is cumbersome for real-time applications. We will look into employing CNAPS using this simple feature vector in the future.

The HLA feature vector also has limitations in real-time implementation. Several steps are required to perform the harmonic matching, which must be tailored to meet real-time criteria. The classification results for the HLA feature space are acceptable; this feature space has the added advantage that it can generate several feature vectors for the same input sequence. It can therefore readily adapt to a multitarget case, where one would like to derive feature vectors for each target present and thus perform multitarget classification. A generalized HLA algorithm should be investigated for performing the multitarget feature extraction. We will use adaptive beamforming to take care of the multitarget issue, but HLA may produce multiple feature vectors, even when an adaptive beamforming technique fails, because the targets are too near each other. Adaptive beamforming is limited in detecting two closely spaced targets, whereas HLA would not have this limit, since it will operate on the received power spectrum and select multiple feature vector examples.

Further analysis will also be performed on the optimum number of features per feature vector for PCA and HLA feature extraction techniques. In the work reported here, the selection for feature vector dimensionality was "ad hoc" at best (primarily driven by the existing HLA algorithm feature space); both classification results and class separability analysis may show that the dimensionality can be reduced further. Preliminary results with the HLA features suggest this to be true for classification of the four-target case. The number of classes included in the analysis should also be extended. More importantly, the four-class problem should be further investigated with data sets recorded in several different environments. Tolerance with respect to the environment is of paramount importance in the evaluation of a feature extraction algorithm for correct classification. We have found that the PSE method is sensitive to environmental conditions in preliminary studies.

Finally, we will investigate features based on wavelet filters, which have been successfully applied in speech recognition and waveform classification [25,26].

References

1. N. Srour and J. Robertson, *Remote Netted Acoustic Detection System*, Army Research Laboratory, ARL-TR-706 (May 1995).
2. G. S. Kang and L. J. Fransen, *Experimentation with an Adaptive Noise Cancellation Filter*, IEEE Trans. Circuits Systems **CAS-34**, No. 7 (July 1987), 753–758.
3. N. Berg, N. Srour, and G. Prado, *Unattended Ground Sensor Performance Prediction and Preliminary System Requirements*, unpublished ARL report (1993).
4. J. A. Robertson and B. Weber, *Artificial Neural Networks for Acoustic Target Recognition*, joint report between ARL and ITT Research (1993).
5. K. Fukunaga, *Introduction to Statistical Pattern Recognition*, second edition, Academic Press, San Diego (1990).
6. J. Mao and A. K. Jain, *Artificial Neural Networks for Feature Extraction and Multivariate Data Projection*, IEEE Trans. Neural Networks **6**, No. 2 (March 1995), 296–317.
7. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, John Wiley and Sons, Inc. (1973).
8. J. V. Ryzin, *Classification and Clustering*, Academic Press, New York (1977).
9. S. Haykin, *Neural Networks, A Comprehensive Foundation*, Macmillan Publishing Company, New Jersey (1994).
10. A. Krogh and J. A. Hertz, *Hebbian Learning of Principal Components*, Parallel Processing in Neural Systems and Computers, Elsevier Science Publishers B. V. (North Holland) (1990), 183–186.
11. E. Oja, *A Simplified Neuron Model as a Principal Component Analyzer*, J. Math. Biol. **15** (1982), 267–273.
12. E. Oja, *Neural Networks, Principal Components, and Subspaces*, Int. J. Neural Syst. **1** (1989), 61–68.
13. *DataBase Mining Workstation*, Reference Manual, HNC, Inc. (1992).
14. A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, New Jersey (1989).
15. K. Hornik and C. M. Kuan, *Convergence Analysis of Local Feature Extraction Algorithm*, Neural Networks **5** (1992), 229–240.
16. J. Rubner and P. Tavan, *A Self Organizing Network for Principal Component Analysis*, Europhys. Lett. **10** (1989), 693–698.
17. M. K. Tsatsanis and G. B. Giannakis, *Principal Component Filter Banks for Optimal Multiresolution Analysis*, IEEE Trans. Sig. Proc. **43**, No. 8 (August 1995).

18. P. D. Welch, *The Use of Fast Fourier Transforms for the Estimation of Power Spectra*, IEEE Trans. Audio-Electroacoustics **AU-15** (June 1970), 70–73.
19. S. Haykin, *Adaptive Filter Theory*, second edition, Prentice Hall, New Jersey (1991).
20. M. Lawrence and A. Peterson, *BrainMaker Professional User's Guide and Reference Manual*, California Scientific Software (1993).
21. D. Hammerstrom, *A VLSI Architecture for High Performance, Low-Cost, On-Chip Learning*, International Joint Conference on Neural Networks, vol. 2, San Diego, CA (1990), 537–544.
22. D. Hammerstrom, *Electronic Neural Network implementation*, Tutorial No. 5, International Joint Conference on Neural Networks, Baltimore, MD (1992).
23. D. Hammerstrom and S. Rahfuss, *Neurocomputing Hardware: Present and Future*, Swedish National Conference on Connectionism, Skovade, Sweden (9–10 September 1992).
24. L. Sirovich and M. Kirby, *Low-Dimensional Procedure for the Characterization of Human Faces*, J. Opt. Soc. Am. **4**, No. 3 (March 1987), 519–524.
25. K. Etemad and R. Chellappa, *Separability Based Tree Structured Basis Selection for Texture Classification*, Proc. First Int. Conf. on Image Processing (November 1994), 441–445.
26. R. E. Learned, W. C. Karl, and A. S. Willsky, *Wavelet Packet Based Transient Signal Classification*, Proc. IEEE Conf. on Time Scale and Time Frequency Analysis (1992), 109–112.

Distribution

| | |
|---|--|
| Admnstr Defns Techl Info Ctr Attn DTIC-OCP 8725 John J Kingman Rd Ste 0944 FT Belvoir VA 22060-6218 | Dept of the Army DCS for Intllgnc Attn DAMI-POB B Allen Washington DC 20310-1001 |
| DIA Central Masint Ofc Attn P Demos Attn CPT M Donofrio 3100 Clarenon Blvd Arlington VA 22201 | Dpty Assist Scy for Rsrch & Techl Attn SARD-TT B Reisman Attn F Milton Attn J Parmentola The Pentagon Washington DC 20310-0103 |
| AATD (ATCOM) Attn AMSAT-R-TU L Sutton FT Eustis VA 22604 | Hdqtrs Dept of the Army Attn DAMO-FDQ MAJ M McGonagle Attn DAMO-FDQ D Schmidt 400 Army Pentagon Washington DC 20310-0460 |
| Army Spc and Strtgc Defns Attn CSSD-TC-C R N Adams PO Box 1500 Huntsville AL 35807-3801 | Hdqtrs Dept of the Army ODSCINT Attn SAMI-POS/MASINT D Bosserman Pentagon Rm 2B459 Washington DC 20301-1001 |
| CECOM NVESD Attn AMSEL-RD-NV-TIS T Anderson Attn AMSEL-RD-NV-SS-GV J Brooks Attn AMSEL-RD-NV-VISP-CR G Klager Attn AMSEL-RD-NV-SS-TPS P Lundy Ste 430 | US Army INSCOM Attn ASEC H J Reckitt 5400 Shawnee Rd Ste 100 Alexandria VA 22312 |
| Attn AMSEL-RD-NV-TISD F Petito Attn AMSEL-RD-NV-TIS J A Rarick Attn AMSEL-RD-NV-TIS-PS D Rehek Attn AMSEL-RD-NV-GSID T Smith Attn AMSEL-RD-NV-SS-GV S Tucker Attn AMSRL-RD-NV-SS-GV R Volpone Ste 430 | INSCOM Dept of the Army Attn B Kitchen Attn E Bielecki Attn M Woodford Attn IAOPS-MA-S B Meyer 8825 Beulah Stret FT Belvoir VA 22060-5246 |
| Attn AMSRL-RD-NV-UAB C Walters FT Belvoir VA 22060 | OSD Attn DUSD/AT C Perkins Pentagon Rm 3E1065 Washington DC 20310-2500 |
| DARPA Attn ASTO T Kooij Attn E Carapezza Attn R Dugan 3701 N Fairfax Arlington VA 22203 | |

Distribution

TARDEC

Hit Avoidance Team
Attn J Ayers
31201 Chicago Rd A301
Warren MI 48093

TASC

Attn D S Bishop
1992 Lewis TGurner
FT Walton FL 32547-1255

US Army AFDD

Ames Rsrch
Attn AMSAT-R-AF D Boxwell
MS N215-1
Attn M/S 219-3 W Mosher
Moffett Field CA 94035

US Army ATCOM

Attn AMSAT-R-TV M Dinning
Attn AMSAT-R-TV G Birocco
FT Eustis VA 23604-5577

US Army CECOM

Attn AMSEL-LC-SM-S3 CPT H Choe
Attn AMSEL-LC-SM-S3 M Dipaola
FT Monmouth NJ 07703-5017

US Army CECOM

Attn AMSEL-RD-NV-CI-NCTR L Stein
FT Monmouth NJ 07703-5206

US Army CECOM/NVESD

Attn AMSEL-RD-NV-RD-IFF M Muller
FT Monmouth NJ 07703

US Army MICOM

Attn AMSMI-RD-MG-IP M W Harper
Bldg 5400
Redstone Arsenal AL 35898

US Army Natick

Attn SSCNY-YB J B Sampson
Natick MA 01760-5020

US Army Rsrch Ofc

Attn AMXRO-PH M Ciftan
PO Box 12211
Research Triangle Park NC 27725

US Army TACOM

Attn AMSTA-TR-S E Shalis
Warren MI 48090

USA ATC

Attn STEC-EN-PS R Joy
Attn STECS-AC-TE-A G Rogers
Aberdeen Proving MD 21005

USA CRREL

Attn CEREL-GP R Detsch
Attn CERL-GP M Moran
72 Lyme Rd
Hanover NH 03755-1290

USA NGIC

Attn IANG-RSG B Grachus
Attn SRA-ELLIOTT C Elliott
220 7th Stret NE
Charlottesville VA 22902

USAE Waterways Experimental Sta

Attn CEWES-SD-R B L Carnes
3909 Halls Ferry Rd
Vicksburg MS 39180-6199

Ofc of Nav Rsrch

Attn R Varley
800 North Quincy
Arlington VA 22217-5660

US Nav Postgraduate Schl

Dept of Physics
Attn K Woehler
Monterey CA 93943

Chicken Little Joint Proj

Attn 46 OG/OGML J A Sledge
104 Cherokee Ave
Eglin AFB FL 32542-5600

Distribution

Lawrence Livermore Natl Lab
Attn M/S L-183 J Baker
7000 East Ave
Livermore CA 94550

NASA Langley Rsrch Ctr
Aeroacoustics Branch
Attn A W Mueller
Fluid Mechanics MS-461
Hampton VA 23681

Johns Hopkins Univ
Applied Physics Lab
Attn A Coon
Johns Hopkins Rd
Laurel MD 20723-6099

Michigan Tech Univ
Keweenaw Rsrch Ctr
Attn J Rogers
1400 Townsend Dr
Houghton MI 49931

Penn State Univ
Appld Rsrch Lab
Attn K C Reichard
Attn D C Swanson
PO Box 30
State College PA 16804

Penn State Univ
Dept of Geosciences
Attn R J Greenfield
Deike Bldg
University Park PA 16801

Univ of Maryland
Ctr for Automtn Rsrch
Attn K-B EOM
Attn R Chellappa
2365 A V Williams Bldg
College Park MD 20742

Univ of Maryland
Elect Engrg Dept
Attn S Shamma
Attn R Liu AV Williams RM 2157
College Park MD 20742-3275

Univ of Michigan
EECS Dept
Attn W J Williams
Systems Div
Ann Arbor MI 48109-2122

Univ of Mississippi
NCPA
Attn H E Bass
University MS 38577

Univ of Texas
Appld Rsrch Lab
Attn N Bedford
Attn M B Bennett
PO Box 8029
Austin TX 78713-8029

Alliant Techsystems
Attn R Braunling
600 2nd Stret NE
Hopkins MN 55343

BBN Sys & Technologies
Attn T Galaitis
70 Fawcett Stret
Cambridge MA 02138

Draper Lab
Image Recognition Sys
Attn M Desai
Attn MS 37 J Bernstein
555 Technology Sq
Cambridge MA 02139

Distribution

| | |
|--------------------------------------|---------------------------------------|
| ENSCO | Northrop Corp |
| Advncd Tech & Planning | Attn T Landers |
| Attn T Cirillo | PO Box 5032 2301 W 120th Stret |
| Attn T Gamble Senior Staff Scientist | Hawthorne CA 90250 |
| 5400 Port Royal Rd | |
| Springfield VA 22151 | |
| Hughes | Northrop Corp |
| Ground Sys Grp | Attn W350/N6 Y-M Chen |
| Attn K Kohnen | PO Box 5032 |
| PO Box 3310 MS E242 | Hawthrone CA 90251-5032 |
| Fullerton CA 92634-3310 | |
| Hughes | Sentech Inc |
| Undersea Sys Div | Attn P Martel |
| Attn M Rakijas | 5 Militia Dr Ste 105 |
| PO Box 3310 MS 604/E242 | Lexington MA 02173 |
| Fullerton CA 92834-3310 | |
| Lockheed Martin Comm Sys | Textron Defns Sys |
| Attn P Walter | Program Dev Tactical Sys |
| Attn S Mui | Attn R J Correia |
| 1 Federal Stret | Attn J Fitzgerald |
| Camden NJ 08102 | Attn R L Steadman |
| Lockheed Sanders Inc | 201 Lowell Stret |
| Attn S W Lang | Wilmington MA 01887 |
| PO Box 2057 | |
| Nashua NH 03061-0868 | |
| Lockheed-Sanders | ARDEC |
| Attn PTP02-A001 D Deadrick | Attn SMCAR-FSF-RM W Donnally |
| PO Box 868 | Attn AMSTA-AR-FSF-R J Heberley |
| Naushua NH 03061-0868 | Sys B95N |
| MIT Lincoln Lab | Picatinney NJ 07801 |
| Attn S4-307B R Lacoss | |
| 244 Wood Stret | |
| Lexington MA 02173-9108 | |
| Mitre Corporation | US Army Rsrch Lab/HRED |
| Attn C Burmaster | Attn AMSRL-HR-SD J Kalb |
| 1820 Dolley Madison | Attn AMSRL-HR-SD T Letowski |
| McLean VA 22102-3481 | Aberdeen Proving MD 21005 |
| | |
| | US Army Rsrch Lab |
| | Attn AMSRL-IS B Broome |
| | Attn AMSRL-IS J D Gantt |
| | Attn AMSRL-IS-C LTC M R Kindl |
| | Attn AMSRL-IS-CI T Mills |
| | Aberdeen Proving Ground MD 21005-5067 |

Distribution

US Army Rsrch Lab/SLAD
Attn AMSRL-SL-EV K Morrison
White Sands Missile Range NM 88002-5513

US Army Rsrch Lab
Attn AMSRL-BE-S H J Avermann
Attn AMSRL-CI-LL Tech Lib (3 copies)
Attn AMSRL-CS-AL-TA Mail & Records
Mgmt
Attn AMSRL-CS-AL-TP Techl Pub (3 copies)

US Army Rsrch Lab (cont'd)
Attn AMSRL-IS-CI D Hillis (8 copies)
Attn AMSRL-SE-EA J Eicke
Attn AMSRL-SE-RU H Khatri
Attn AMSRL-SE-SA N Srour (8 copies)
Attn AMSRL-SE-SA R Wellman
(8 copies)
Adelphi MD 20783-1197

| REPORT DOCUMENTATION PAGE | | | <i>Form Approved OMB No. 0704-0188</i> |
|---|--|--|--|
| <p>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.</p> | | | |
| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED | |
| | January 1997 | Interim; September 1995 to May 1996 | |
| 4. TITLE AND SUBTITLE | | 5. FUNDING NUMBERS | |
| Acoustic Feature Extraction for a Neural Network Classifier | | DA PR: AH16 PE: 62120 | |
| 6. AUTHOR(S) | | | |
| Mark C. Wellman, Nassy Srour, and David B. Hillis | | | |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) | | 8. PERFORMING ORGANIZATION REPORT NUMBER | |
| U.S. Army Research Laboratory Attn: AMSRL-SE-RM 2800 Powder Mill Road Adelphi, MD 20783-1197 | | ARL-TR-1166 | |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | | 10. SPONSORING/MONITORING AGENCY REPORT NUMBER | |
| U.S. Army Research Laboratory 2800 Powder Mill Road Adelphi, MD 20783-1197 | | | |
| 11. SUPPLEMENTARY NOTES ARL PR: 7NE4TT AMS code: 622120.H1611 | | | |
| 12a. DISTRIBUTION/AVAILABILITY STATEMENT | | 12b. DISTRIBUTION CODE | |
| Approved for public release; distribution unlimited. | | | |
| 13. ABSTRACT (Maximum 200 words) | | | |
| <p>Artificial neural networks can perform reliable classification of ground vehicles based solely on their acoustic signatures, if robust features can be identified. We present feature extraction and classification results using simple power spectrum estimates, harmonic line association, and principal component analysis. Algorithm implementation and performance analysis of each feature extraction method are discussed. Also given are preliminary evaluation results of a VLSI (very-large-scale integration) device dedicated to neural network implementation.</p> | | | |
| 14. SUBJECT TERMS | | 15. NUMBER OF PAGES | |
| Acoustics, feature extraction | | 26 | |
| | | 16. PRICE CODE | |
| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
| Unclassified | Unclassified | Unclassified | UL |